11-2013

DNA Barcoding Distinguishes Pest Species of the Black Fly Genus *Cnephia* (Diptera: Simuliidae)

I. M. Confitti
University of Toronto

K. P. Pruess
University of Nebraska-Lincoln

A. Cywinska
Ingenomics, Inc.

T. O. Powers
University of Nebraska-Lincoln

D. C. Currie
University of Toronto and Royal Ontario Museum, dc.currie@utoronto.ca

DNA Barcoding Distinguishes Pest Species of the Black Fly Genus *Cnephia* (Diptera: Simuliidae)

I. M. CONFLITTI,[1,2] K. P. PRUESS,[3] A. CYWINSKA,[4] T. O. POWERS,[5] AND D. C. CURRIE[1,2,6]

J. Med. Entomol. 50(6): 1250–1260 (2013); DOI: http://dx.doi.org/10.1603/ME13063

ABSTRACT Accurate species identification is essential for cost-effective pest control strategies. We tested the utility of COI barcodes for identifying members of the black fly genus *Cnephia* Enderlein (Diptera: Simuliidae). Our efforts focus on four Nearctic *Cnephia* species—*Cnephia dacotensis* (Dyar & Shannon), *Cnephia eremities* Shewell, *Cnephia ornithophilia* (Davies, Peterson & Wood), and *Cnephia pecuarum* (Riley)—the latter two being current or potential targets of biological control programs. We also analyzed one Palearctic species, *Cnephia pallipes* (Fries). Although *Cnephia* adults can be identified anatomically to species, control programs target the larval stage, which is difficult or impossible to distinguish morphologically. By using neighbor-joining, maximum parsimony, and Bayesian methods, we found that COI barcodes successfully identified three Nearctic *Cnephia* species, but not *C. pecuarum*. The Palearctic *C. pallipes* was also successfully identified. Despite nonmonophyly of *C. pecuarum*, we show that data from COI barcoding, in combination with geographical and ecological information, can be used to distinguish all four Nearctic species. Finally, we discussed 1) possible reasons for paraphyly in *C. pecuarum*, 2) topological concordance to previously reported chromosomal dendrograms, and 3) evolution of diverse feeding strategies within the genus *Cnephia*.

KEY WORDS biological control, COI gene, mitochondrial DNA, monophyly, paraphyly

Species identification is important for pest control programs that must consider both the costs and efficacy of pest mitigation efforts. One group of organisms that have been proven especially difficult to identify are black flies—a worldwide family of nematocerous Diptera that is infamous for the bloodsucking habits of females and transmission of the causal agents of parasitic diseases to birds and mammals. Members of the family are structurally homogenous, making routine "hard parts" identification difficult or even impossible in many instances. Moreover, cytological studies reveal that many nominal species consist of two or more reproductively isolated (but structurally identical) sibling species, or "cytospecies." Cytological identification is typically possible only through band-by-band analyses of the giant polytene chromosomes of larval black flies, meaning that other life-history stages are often impossible to identify. The emergence of DNA barcoding—the use of a short DNA sequence to identify organisms to a particular species—offers the prospect that hitherto intractable species or life-history stages can now be identified (Hebert et al. 2003a,b). Preliminary research reveals that DNA barcoding holds sufficient promise for black fly identification to warrant further study (Day et al. 2010, Hunter et al. 2008, Ilmonen et al. 2009, Pramual et al. 2011, Rivera and Currie 2009).

Cnephia Enderlein is a small genus of black flies with eight species distributed throughout the Holarctic Region (Adler and Crosskey 2012). Four species occur in the Nearctic Region, including the northern Holarctic species *Cnephia eremities* Shewell, and three endemic species: *Cnephia dacotensis* (Dyar & Shannon), *Cnephia ornithophilia* (Davies, Peterson & Wood), and *Cnephia pecuarum* (Riley). Although represented by few species, the North American *Cnephia* exhibit the full range of feeding habits of female black flies. Two species are obligately autogenous (*C. eremites* and *C. dacotensis*), one is ornithophilic (*C. ornithophilia*), and one is mammalophilic (*C. pecuarum*). *C. ornithophilia* seeks blood from a wide variety of avian hosts and is a vector of *Leucocytozoon* Ziemann protozoa to woodland birds (Fallis and Bennett 1961, 1962; Khan and Fallis 1970). They pose a particular threat to the endangered Attwater's Prairie-Chicken (*Tympanuchus cupido attwateri* Bendire) through the bloodsucking activity and possible transmission of *Leucocytozoon* (Adler et al. 2007). Control measures against *C. ornithophilia* were suggested as a possibility, if warranted, in streams and rivers in the vicinity of the Attwater

[1] Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks St., Toronto, ON M5S 3B2, Canada.
[2] Department of Natural History, Royal Ontario Museum, 100 Queen's Park, Toronto, ON M5S 2C6, Canada.
[3] Department of Entomology, University of Nebraska, 103 Entomology Hall, Lincoln, NE 68583.
[4] Ingenomics Inc., 3164 Candela Dr., Mississauga, ON L5A 2T8, Canada.
[5] Department of Plant Pathology, University of Nebraska, 406 Plant Science Hall, Lincoln, NE 68583.
[6] Corresponding author, e-mail: dc.currie@utoronto.ca.

Prairie-Chicken National Wildlife Refuge in Colorado Co., Texas (Adler et al. 2007). *C. pecuarum*—the southern buffalo gnat of the Mississippi River Valley—is among the most notorious bloodsucking pests of humans and other mammals in North America (Adler et al. 2004). Massive mortality of livestock was attributed to this species during the American Civil War, when levees on the Mississippi River deteriorated and river water flowed onto the brushy alluvial zone, creating ideal breeding sites (Riley 1887). Mortality was probably the result of acute toxemia and anaphylactic shock caused by the injection of saliva during blood feeding (Adler and McCreadie 2002). Simuliotoxicosis on such a scale is now rare; however, occasional outbreaks of *C. pecuarum* make them the target of control programs in localized areas throughout their range (Adler et al. 2004).

Although *Cnephia* adults can be identified anatomically to species, the immature stages of most species are difficult or impossible to distinguish morphologically. This is an impediment to biological studies that require accurate identification of larvae and pupae; and it is especially problematical for control programs, which focus mainly on larval control. The foremost biological control agent for black flies—the bacterium *Bacillus thuringiensis* variety *israelensis* (Bti)—must be applied to breeding sites in a timely and judicious fashion to increase efficacy and reduce costs (Molloy et al. 1981). As the two pest species (*C. pecuarum* and *C. ornithophilia*) are sympatric, it is impossible to know which breeding sites to target unless 1) relatively mature larvae are available for cytotyping, or 2) pupae have reached the pharate adult stage. Neither situation is desirable, as there would be little time to implement a larvaciding program before pestiferous adults have emerged. DNA barcoding, if effective, would permit identification at a much earlier stage of development, as immature larvae or even eggs could be screened from prospective breeding sites. In this study, we test the diagnostic utility of a 615-bp fragment of the cytochrome c oxidase subunit I (COI) gene as a DNA barcode for the Nearctic species of *Cnephia.*

Materials and Methods

Taxon Sampling, DNA Extraction, Amplification, and Sequencing. Ingroup sampling consisted of 90 individuals representing five species within the genus *Cnephia* (four Nearctic species and one exclusively Palearctic species, *C. pallipes* (Fries); Table 1). We included multiple individuals per species and, where possible, sampled from various localities throughout their geographic range. Eight exemplars of *Stegopterna decafilis* Rubtsov were chosen as the outgroup based on morphological evidence (Adler et al. 2004) and sample availability. Outgroup sequences were obtained from the Barcode of Life Data Systems Website (http://www.boldsystems.org).

Genomic DNA was extracted from the thorax and abdomen of larvae or pupae preserved in 90–95% ethanol or from 1 to 3 legs of adult specimens. For larval specimens of *C. dacotensis* and *C. ornithophilia*, which are morphologically identical, identifications were based either on distributional data (i.e., specimens were assigned to a species if it occurred decisively within that species' range and was also clearly separated from the known ranges of all other species) or on cytologically screened specimens from the same sites. For pupae, identifications were based either by using the same distributional criteria as described for larvae, or on fully developed pupae (i.e., pharate adults) whose genitalia could be examined. Vouchers for specimens are held in the entomological collections of the University of Nebraska State Museum and Royal Ontario Museum. Extractions were performed by using the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma–Aldrich Co., St. Louis, MO) or the Gentra Puregene Kit (Qiagen, Germantown, MD), following manufacturer's instructions. Standard DNA barcoding primers (LCO1490 and HCO2198, Folmer et al. 1994) were used to target a 658-bp region of the mitochondrial COI gene. Each 25 μl polymerase chain reaction (PCR) consisted of 0.92× PCR buffer pH 8.3 (10 mM Tris-HCl pH 8.3, 50 mM KCl, and 0.01% NP-40), 2.6 mM $MgCl_2$, 200 μM of each dNTP, 0.3 μM of each primer, 1 U *Taq*DNA polymerase, 1–5 μl template DNA, and the remaining volume of dH_2O. PCR cycling parameters were as follows: 1 min at 95°C; 35 cycles of 1 min at 94°C, 1 min at 55°C, and 1.5 min at 72°C; and finally 7 min at 72°C. Samples were sequenced with the ABI Big Dye 3.1 Terminator Kit (Applied Biosystems Inc., Carlsbad, CA). Sequencing primers were the same as those used for PCR. Sequenced products were run on ABI 3730 DNA Analyzers (Applied Biosystems Inc.). Table 1 contains GenBank accession numbers for each DNA sequence obtained.

Data Analysis. COI sequences were automatically aligned using Sequencher v4.5 (Gene Codes Corp., Ann Arbor, MI). The resulting alignments were further inspected visually in BioEdit Sequence Alignment Editor v7.0.9.0 (Hall 1999). Pairwise sequence divergence within and between species was calculated in PAUP* v4.0b10 (Swofford 2003) using the Kimura 2-parameter (K2P) model (Kimura 1980). Summary statistics of intraspecific and interspecific distances (i.e., minimum, maximum, and mean) as well as histograms were produced using Microsoft Excel 2007 (Microsoft Corp., Redmond, WA).

Phylogenetic trees were constructed using neighbor-joining (NJ), maximum parsimony (MP), and Bayesian methods. NJ analysis of K2P distances (Saitou and Nei 1987) was conducted in PAUP*. Clade support was estimated using 10,000 bootstrap replicates (faststep search; Felsenstein 1985). MP trees were also constructed in PAUP*. Constant and uninformative characters were excluded from the analysis and trees were rooted with a monophyletic outgroup. All characters were considered unweighted and unordered. Heuristic searches were performed using 1,000 replicates of random sequence addition and TBR (tree bisection and reconnection) branch swapping. Clade support was estimated from the entire data set

Table 1. List of samples used in this study with corresponding collection information and GenBank accession numbers

Taxon	Code	Locality	Site	Date	GenBank no.
Stegopterna decafilis	YT-1	Yukon Territory, Canada	Dempster Highway, 88 km north of Highway 2	N/A	KC294446
S. decafilis	YT-2	Yukon Territory, Canada	Dempster Highway, 88 km north of Highway 2	N/A	KC294447
S. decafilis	YT-3	Yukon Territory, Canada	Dempster Highway, 88 km north of Highway 2	N/A	KC294448
S. decafilis	YT-4	Yukon Territory, Canada	Dempster Highway, 88 km north of Highway 2	N/A	KC294449
S. decafilis	YT-5	Yukon Territory, Canada	Dempster Highway, 88 km north of Highway 2	N/A	KC294450
S. decafilis	YT-6	Yukon Territory, Canada	Dempster Highway, 75.5 km north of Highway 2	N/A	KC294451
S. decafilis	YT-7	Yukon Territory, Canada	Dempster Highway, 75.5 km north of Highway 2	N/A	KC294452
S. decafilis	YT-8	Yukon Territory, Canada	Dempster Highway, 75.5 km north of Highway 2	N/A	KC294453
Cnephia dacotensis	AB-1	Alberta, Canada	Rosa Creek, east of Elk Island	4 May 1999	KC294454
C. dacotensis	AB-2	Alberta, Canada	Rosa Creek, east of Elk Island	4 May 1999	KC294455
C. dacotensis	AB-3	Alberta, Canada	Rosa Creek, east of Elk Island	4 May 1999	KC294456
C. dacotensis	AB-4	Alberta, Canada	Rosa Creek, east of Elk Island	4 May 1999	KC294457
C. dacotensis	KS-1	Kansas, USA	North Fork Frog Creek, Coffey County	2 April 2003	KC294458
C. dacotensis	KS-2	Kansas, USA	North Fork Frog Creek, Coffey County	2 April 2003	KC294459
C. dacotensis	KS-3	Kansas, USA	North Fork Frog Creek, Coffey County	2 April 2003	KC294460
C. dacotensis	MB-1	Manitoba, Canada	Highway 16 West, Neepawa	14 May 2005	KC294461
C. dacotensis	MB-2	Manitoba, Canada	Highway 16 West, west of Neepawa	14 May 2005	KC294462
C. dacotensis	MB-3	Manitoba, Canada	Highway 16 West, west of Neepawa	14 May 2005	KC294463
C. dacotensis	MB-4	Manitoba, Canada	Highway 366, Duck Mountain Provincial Park	15 May 2005	KC294464
C. dacotensis	MB-5	Manitoba, Canada	Highway 366, 5.5 km south of Baldy Mountain Road	15 May 2005	KC294465
C. dacotensis	MB-6	Manitoba, Canada	Trans-Canada Highway 1, between Brandon & Routhwaite	18 May 2005	KC294466
C. dacotensis	MB-7	Manitoba, Canada	Trans-Canada Highway 1, between Brandon & Routhwaite	18 May 2005	KC294467
C. dacotensis	MB-8	Manitoba, Canada	Trans-Canada Highway 1, between Brandon & Routhwaite	18 May 2005	KC294468
C. dacotensis	MN-1	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294469
C. dacotensis	MN-2	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294470
C. dacotensis	MN-3	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294471
C. dacotensis	MN-4	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294472
C. dacotensis	MN-5	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294473
C. dacotensis	MN-6	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294474
C. dacotensis	MN-7	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294475
C. dacotensis	MN-8	Minnesota, USA	Rush Creek, Hennepin County	26 April 2007	KC294476
C. dacotensis	NE-1	Nebraska, USA	Yankee Hill Road & South 70th Street, Lancaster County	4 April 2004	KC294477
C. dacotensis	NE-2	Nebraska, USA	Yankee Hill Road & South 70th Street, Lancaster County	4 April 2004	KC294478
C. dacotensis	NE-3	Nebraska, USA	Yankee Hill Road & South 70th Street, Lancaster County	4 April 2004	KC294479
C. dacotensis	ON-1	Ontario, Canada	Algonquin Provincial Park, Whitney	9 Aug. 2005	KC294480
C. dacotensis	ON-2	Ontario, Canada	Algonquin Provincial Park, Whitney	9 Aug. 2005	KC294481
C. dacotensis	ON-3	Ontario, Canada	Algonquin Provincial Park, Whitney	9 Aug. 2005	KC294482
C. dacotensis	ON-4	Ontario, Canada	Algonquin Provincial Park, Whitney	9 Aug. 2005	KC294483
C. dacotensis	ON-5	Ontario, Canada	Algonquin Provincial Park, Whitney	9 Aug. 2005	KC294484
C. dacotensis	ON-6	Ontario, Canada	Algonquin Provincial Park, Whitney	6 May 1999	KC294485
C. dacotensis	ON-7	Ontario, Canada	Algonquin Provincial Park, Whitney	6 May 1999	KC294486
C. dacotensis	ON-8	Ontario, Canada	Algonquin Provincial Park, Whitney	26 May 2004	KC294487
C. dacotensis	ON-9	Ontario, Canada	Algonquin Provincial Park, Whitney	6 May 1999	KC294488
C. dacotensis	PA-1	Pennsylvania, USA	White Heron Lake Outlet, Monroe County	28 Mar. 2003	KC294489
C. dacotensis	PA-2	Pennsylvania, USA	White Heron Lake Outlet, Monroe County	28 Mar. 2003	KC294490
C. dacotensis	PA-3	Pennsylvania, USA	White Heron Lake Outlet, Monroe County	28 Mar. 2003	KC294491
C. dacotensis	PA-4	Pennsylvania, USA	White Heron Lake Outlet, Monroe County	28 Mar. 2003	KC294492
C. eremites	FN-1	Finland	3 km south of Virtaniemi	20 June 2005	KC294493
C. eremites	FN-2	Finland	3 km south of Virtaniemi	20 June 2005	KC294494
C. eremites	FN-3	Finland	3 km south of Virtaniemi	20 June 2005	KC294495
C. eremites	NU-1	Nunavut, Canada	Prince River, 14 km north east of Baker Lake	15 July 2003	KC294496
C. eremites	NU-2	Nunavut, Canada	outflow of Little Meliadine River, Rankin Inlet	19 July 2003	KC294497
C. eremites	NU-3	Nunavut, Canada	outflow of Little Meliadine River, Rankin Inlet	19 July 2003	KC294498
C. eremites	NU-4	Nunavut, Canada	Prince River, 14 km north east of Baker Lake	15 July 2003	KC294499
C. eremites	NU-5	Nunavut, Canada	Prince River, 14 km north east of Baker Lake	15 July 2003	KC294500
C. eremites	SW-1	Sweden	Skattan	15 June 2003	KC294501
C. eremites	SW-2	Sweden	Skattan	15 June 2003	KC294502
C. eremites	SW-3	Sweden	Skattan	15 June 2003	KC294503
C. ornithophilia	FL-1	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294504
C. ornithophilia	FL-2	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294505
C. ornithophilia	FL-3	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294506
C. ornithophilia	FL-4	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294507
C. ornithophilia	FL-5	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294508
C. ornithophilia	FL-6	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294509

Continued on following page

Table 1. Continued

Taxon	Code	Locality	Site	Date	GenBank no.
C. ornithophilia	FL-7	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294510
C. ornithophilia	FL-8	Florida, USA	Little Alapaha River, Hamilton County	4 Feb. 2003	KC294511
C. ornithophilia	KS-1	Kansas, USA	North Fork Frog Creek, Coffey County	2 April 2003	KC294512
C. ornithophilia	KS-2	Kansas, USA	North Fork Frog Creek, Coffey County	2 April 2003	KC294513
C. ornithophilia	MD-1	Maryland, USA	Cash Creek, Prince George's County	1 May 2003	KC294514
C. ornithophilia	MD-2	Maryland, USA	Cash Creek, Prince George's County	1 May 2003	KC294515
C. ornithophilia	NL-1	Newfoundland, Canada	Hughes Pond Outlet, St. John's	3 Mar. 2003	KC294516
C. ornithophilia	NL-2	Newfoundland, Canada	Hughes Pond Outlet, St. John's	3 Mar. 2003	KC294517
C. ornithophilia	OK-1	Oklahoma, USA	Sand Creek, Osage County	1 April 2003	KC294518
C. ornithophilia	OK-2	Oklahoma, USA	Sand Creek, Osage County	1 April 2003	KC294519
C. ornithophilia	OK-3	Oklahoma, USA	Sand Creek, Osage County	1 April 2003	KC294520
C. ornithophilia	OK-4	Oklahoma, USA	Sand Creek, Osage County	1 April 2003	KC294521
C. ornithophilia	ON-1	Ontario, Canada	Algonquin Provincial Park, Whitney	6 May 1999	KC294522
C. ornithophilia	PA-1	Pennsylvania, USA	White Heron Lake Outlet, Monroe County	28 Mar. 2003	KC294523
C. ornithophilia	SC-1	South Carolina, USA	Woods Bay State Park, Sumter County	16 Mar. 1998	KC294524
C. ornithophilia	SC-2	South Carolina, USA	Woods Bay State Park, Sumter County	16 Mar. 1998	KC294525
C. ornithophilia	SC-3	South Carolina, USA	Woods Bay State Park, Sumter County	16 Mar. 1998	KC294526
C. ornithophilia	TX-1	Texas, USA	Sabine River, Wood County	25 Jan. 1993	KC294527
C. ornithophilia	TX-2	Texas, USA	Sabine River, Wood County	5 Jan. 2003	KC294528
C. ornithophilia	TX-3	Texas, USA	Sabine River, Wood County	5 Jan. 2003	KC294529
C. pallipes	FN-1	Finland	Marrakoski	18 June 2005	KC294530
C. pallipes	FN-2	Finland	Marrakoski	18 June 2005	KC294531
C. pallipes	SW-1	Sweden	Svanstein	20 June 2003	KC294532
C. pallipes	SW-2	Sweden	Svanstein	20 June 2003	KC294533
C. pallipes	SW-3	Sweden	Råne River	22 June 2003	KC294534
C. pallipes	SW-4	Sweden	Råne River	22 June 2003	KC294535
C. pallipes	SW-5	Sweden	Råne River	22 June 2003	KC294536
C. pecuarum	TX-1	Texas, USA	Wood County	15 Jan 1993	KC294537
C. pecuarum	TX-2	Texas, USA	Sabine River, Wood County	25 Jan 1993	KC294538
C. pecuarum	TX-3	Texas, USA	Sabine River, Wood County	28 Jan 1993	KC294539
C. pecuarum	TX-4	Texas, USA	Sabine River, Wood County	28 Jan 1993	KC294540
C. pecuarum	TX-5	Texas, USA	Sabine River, Wood County	28 Jan 1993	KC294541
C. pecuarum	TX-6	Texas, USA	Sabine River, Wood County	25 Jan 1993	KC294542
C. pecuarum	TX-7	Texas, USA	Sabine River, Wood County	5 Jan. 2003	KC294543

N/A, not available.

(constant and uninformative characters included) using 10,000 bootstrap replicates (faststep search; Felsenstein 1985).

In terms of Bayesian inference, the Akaike Information Criterion (Posada and Buckley 2004) within MrModelTest v2.2 (Nylander 2004) was used to determine the best fitting evolutionary model for the data set. Bayesian analysis was performed with MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001, Ronquist and Huelsenbeck 2003) by Markov Chain Monte Carlo sampling for 5.0×10^6 generations. Because MrBayes only allows for a single outgroup individual, *Stegopterna decafilis* YT-4 was randomly chosen for this position. The analysis consisted of two simultaneous runs with random starting trees and six Markov chains sampled every 100 generations. The first 25% of parameter estimates was discarded as burn-in. Then, stationarity was confirmed by plotting -ln likelihood (-lnL) scores against generation time. A 50% majority-rule consensus tree was constructed from 37,500 post burn-in trees.

Results

We obtained 90 COI sequences, 615 bp in length, for five *Cnephia* species (Table 1). No insertions, deletions, or stop codons were found, indicating that all sequences constitute functional mitochondrial products (Funk and Omland 2003). Consistent with results from other insect studies, our mtDNA data set showed strong A + T bias in base composition (A: 27.8%, T: 37.5%, G: 17.0%, C: 17.7%; A + T = 65.3%; Clary and Wolstenholme 1985, Nardi et al. 2001).

Within the ingroup, intraspecific pairwise sequence divergence based on the K2P model ranged from 0.00 to 7.84%, with a mean of 1.63% (Fig. 1). The highest divergence value was found in *C. ornithophilia*, whereas the lowest was present in all *Cnephia* species, except *C. pecuarum.* Interspecific K2P distances ranged from 1.49 to 13.35% (mean = 8.47; Fig. 1). Therefore, identical COI sequences were shared within, but not between, *Cnephia* species. The slight overlap between distances (Fig. 1) resulted mainly from high intraspecific divergence values within *C. ornithophilia* and low interspecific distances for both *C. dacotensis* and *C. pecuarum*, as compared with all other *Cnephia* species. Pairwise distances between ingroup and outgroup taxa varied from 13.53 to 16.80%, with a mean of 15.52%.

The NJ tree based on K2P distances showed monophyly of the genus *Cnephia* (bootstrap support [BSS] = 100%; Fig. 2). All species were monophyletic, except *C. pecuarum.* In each case, species monophyly received either moderate or high BSS (*C. dacotensis:* BSS = 74%; *C. eremites:* BSS = 100%; *C. ornithophilia:* BSS = 89%; and *C. pallipes:* BSS = 100%). *C. pecuarum* could not

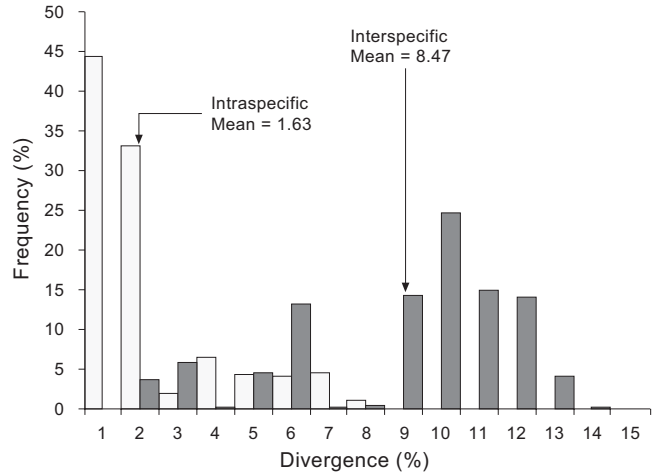

Fig. 1. Frequency distribution of intraspecific (light gray) and interspecific (dark gray) K2P distances for five *Cnephia* species.

be distinguished using the barcoding gene. Sequences of this species clustered with those of *C. dacotensis*, rather than with conspecific barcodes.

Of 615 COI nucleotides, 467 were excluded from MP analysis, 458 of which were constant and nine uninformative. The remaining 148 informative molecular characters were analyzed under the parsimony criterion, yielding 300 equally parsimonious trees, 367 steps in length. The consistency index, homoplasy index, retention index, and rescaled consistency index were 0.5313, 0.4687, 0.9381, and 0.4985, respectively. The parsimony strict consensus topology showed considerably less resolution than the NJ tree. However, both analyses produced consistent results, except for the position of *C. pallipes*. Given the similarity between MP and NJ topologies, BSS for MP analysis is presented on the NJ tree (Fig. 2).

For Bayesian analysis, COI sequences were assigned the general time reversible model of evolution with invariant sites (I) and gamma-distributed rate heterogeneity (G). Our Bayesian phylogeny of relationships among *Cnephia* species is presented in Fig. 3. In comparing NJ, MP, and Bayesian topologies, all three methods produced similar results. However, differences in some poorly supported intraspecific relationships are apparent. The three methods of analysis also recovered differences in the position of *Cnephia* species. For the most part, nodes associated with conflicting interspecific relationships are poorly supported. Despite these topological differences, the following conclusions can be drawn.

First, all species within the genus *Cnephia*, except *C. pecuarum*, were monophyletic (Figs. 2 and 3). Support for monophyly ranged from significant in *C. eremites* and *C. pallipes* (NJ and MP BSS = 100%; Bayesian posterior probabilities [BPP] = 1.00) to moderate, but variable across methods, in *C. dacotensis* (NJ BSS = 74%; MP BSS = 66%; BPP = 0.88) and *C. ornithophilia* (NJ BSS = 89%; MP BSS = 68%; BPP = 0.97). Support for nodes associated with nonmonophyly of *C. pecuarum* is low in all analyses. Second, the position of *C. pallipes* remains unresolved (Figs. 2 and 3). NJ and Bayesian methods place *C. pallipes* at the base of the ingroup. In contrast, this species forms the sister group to all *Cnephia* taxa, except *C. ornithophilia*, in the MP tree. Significant support associated with the position of *C. pallipes* is lacking in all three topologies. Third, depending on branching of *C. pallipes*, *C. ornithophilia* is either positioned at the base of the tree (MP analysis) or as the sister group to the clade containing *C. dacotensis*, *C. eremites*, and *C. pecuarum* (NJ and Bayesian analyses; Figs. 2 and 3). Finally, the relationships among *C. dacotensis*, *C. eremites*, and *C. pecuarum* depend on the placement of *C. pecuarum* individuals (Figs. 2 and 3). All three analysis methods place *C. pecuarum* TX-3 as sister to *C. dacotensis.* The remaining six *C. pecuarum* individuals are either associated with TX-3 + *C. dacotensis* (NJ and MP tree) or split into two groups and distributed throughout the clade containing *C. dacotensis*, *C. eremites*, and *C. pecuarum* samples (Bayesian tree).

Discussion

Utility of COI Gene for Barcoding *Cnephia* Species. DNA barcoding successfully identified three of four Nearctic *Cnephia* species, the exception being the apparent paraphyly of *C. pecuarum* (Figs. 2 and 3). The Palearctic species, *C. pallipes*, was also successfully identified based on DNA barcodes. The apparent nonmonophyly of *C. pecuarum* notwithstanding, data from COI barcoding, in combination with geographical and ecological information, can be used to distinguish all four Nearctic species.

C. pecuarum exhibits a unique distribution among black flies, in that it breeds exclusively in the Mississippi River Valley, from Illinois and Indiana south to the Gulf Coast (Adler et al. 2004). This species is broadly sympatric with only *C. ornithophilia*, from which it can be distinguished unambiguously using

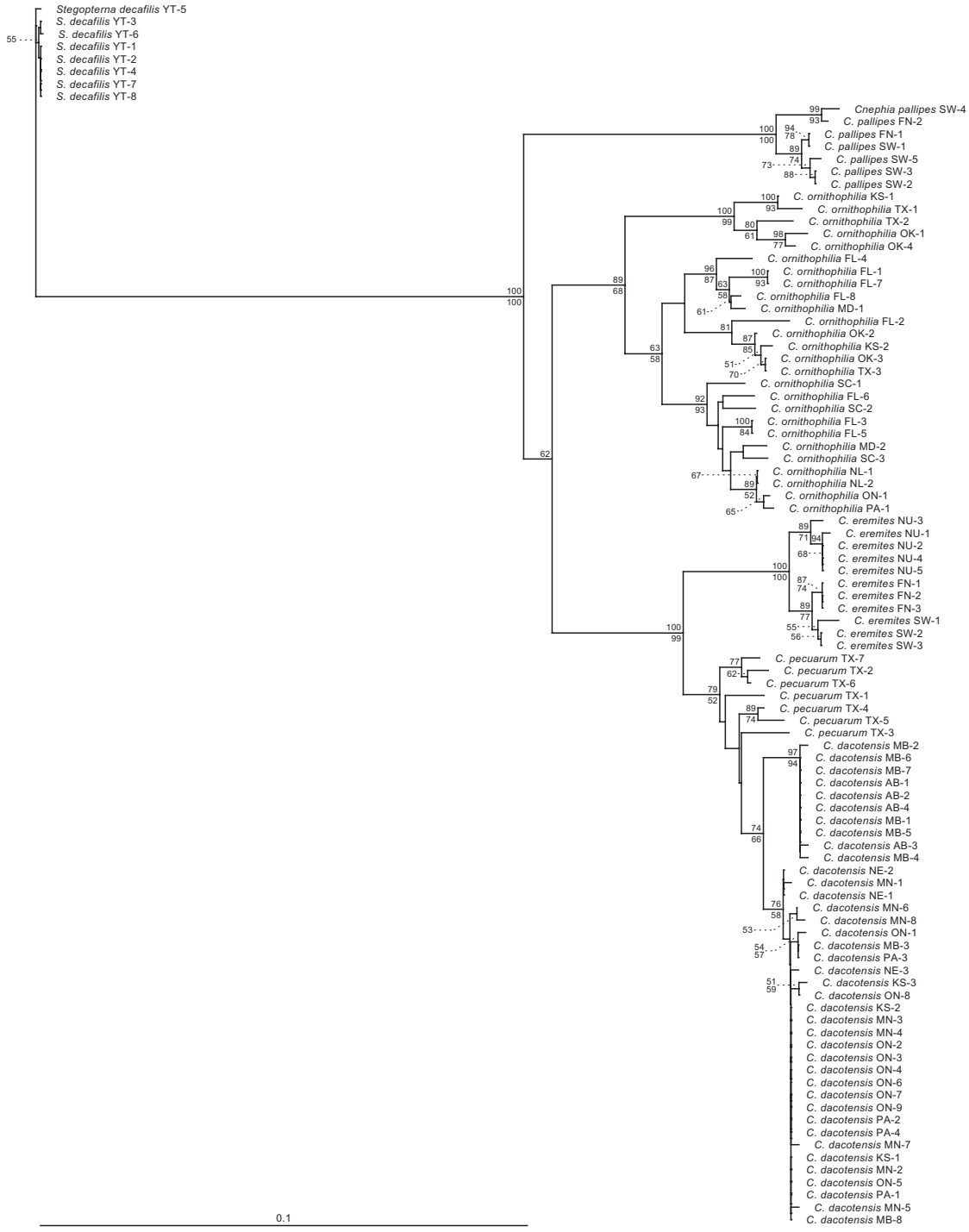

Fig. 2. Neighbor-joining tree generated using K2P distances. Analysis is based on 90 *Cnephia* barcode sequences, 615 bp in length. Numbers above and below the nodes indicate bootstrap support (>50%) for neighbor-joining and maximum parsimony analyses, respectively.

barcoding data (Figs. 2–4). Although the northern distribution of *C. pecuarum* is extended into the southernmost distribution for *C. dacotensis* (Fig. 4), these two species have never been collected together. The former breeds in large-sized rivers, whereas the latter is found in smaller-sized productive streams, such as lake and pond outlets associated with pastures and feedlots (Adler et al. 2004). Accordingly,

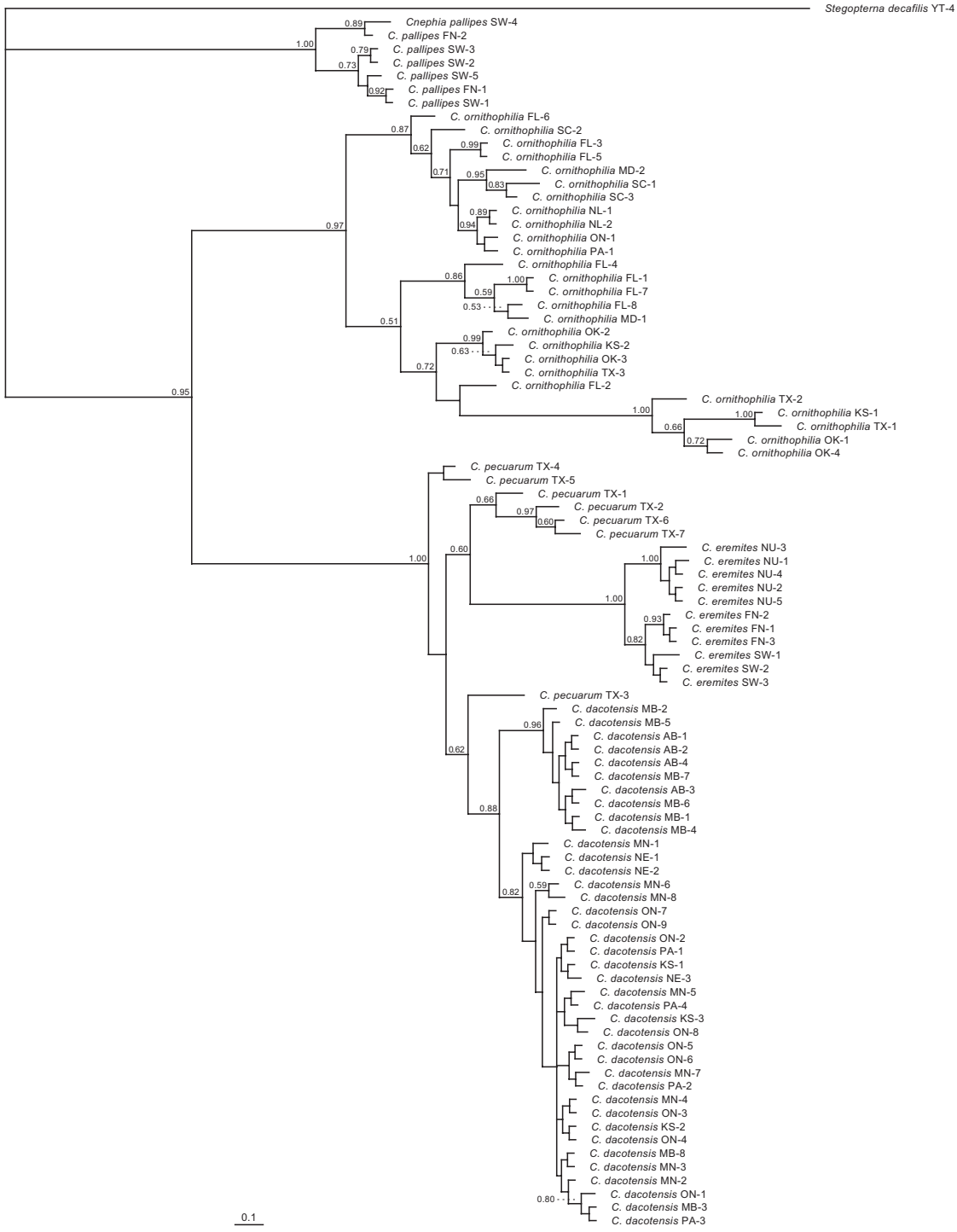

Fig. 3. Phylogenetic relationships among five *Cnephia* species recovered with Bayesian methods. The 50% majority-rule consensus tree sampled from the posterior distribution is presented (-lnL: 2902.12). Numbers above the nodes indicate Bayesian posterior probabilities (>0.50).

they are unlikely misidentified if basic locality information is associated with specimens. However, even if such data are lacking, the monophyly of *C. dacotensis* relative to a paraphyletic *C. pecuarum* ensures that specimens can be assigned confidently to species.

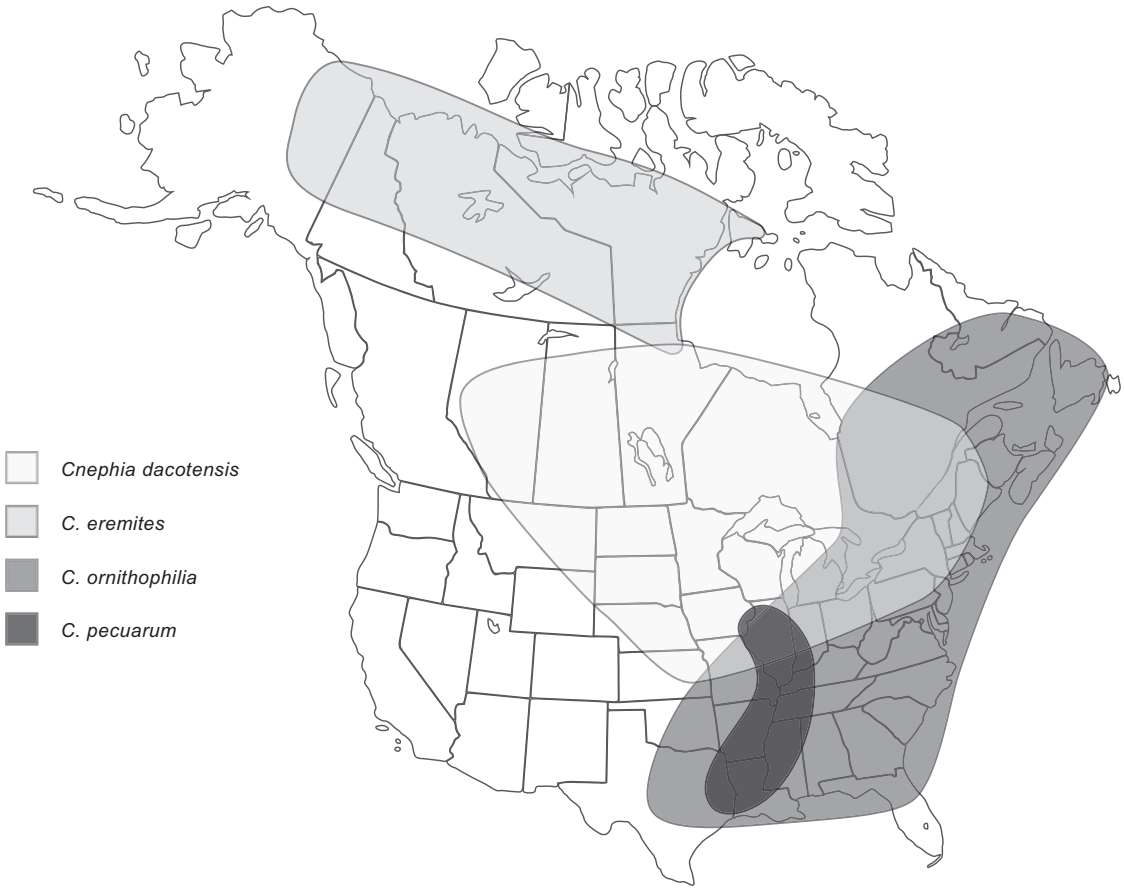

Fig. 4. Distribution of *Cnephia* species in North America.

Possible Reasons for Paraphyly in *C. pecuarum*. Funk and Omland (2003) recognized five sources of mitochondrial DNA failure to detect species-level monophyly—gene paralogy, imperfect taxonomy, inadequate genetic information, incomplete lineage sorting, and/or introgressive hybridization. We address each process to interpret paraphyly in *C. pecuarum*.

The presence of undetected paralogous gene sequences is not likely the cause of nonmonophyly. No insertions, deletions, or stop codons were encountered in our data set, indicating that all sequences constitute functional mitochondrial products (Funk and Omland 2003). We also reject imperfect taxonomy as the cause of nonmonophyly. Considerable evidence supports species status of *C. pecuarum*, including adult morphology, chromosomal banding pattern, feeding mode, and geographic distribution (Adler et al. 2004). Furthermore, misidentification of immature *C. pecuarum* specimens as either *C. dacotensis* or *C. eremites* is not likely, given our sampling localities (Table 1) and each species distribution (Fig. 4). Our analysis is consistent with inadequate genetic information in the COI gene. Evidence includes lack of support for bifurcations associated with paraphyly (Fig. 2, BSS $<$ 50%; Fig. 3, BPP $\leq$ 0.62) and instability of the position of *C. pecuarum* individuals across analysis methods (Figs. 2 and 3). In terms of incomplete lineage sorting, coalescent-based species delimitation using multilocus data may be necessary to rule out this explanation of *C. pecuarum* paraphyly (Fujita et al. 2012). Finally, introgressive hybridization could not account for the patterns observed here. Neither *C. dacotensis* nor *C. eremites* share identical COI sequences with *C. pecuarum*. Moreover, the allopatric distribution of *C. pecuarum* relative to these two species (Fig. 4) would not facilitate hybridization.

We conclude that incomplete lineage sorting and/or inadequate genetic information is likely responsible for paraphyly in *C. pecuarum*. More intensive sampling—in terms of the number and type of genes analyzed (i.e., mitochondrial vs. nuclear DNA), as well as the quantity and distribution of collected *C. pecuarum* individuals—may help resolve monophyly of *C. pecuarum*.

Evolutionary Insights. *Cnephia* is an enigmatic genus whose members differ markedly from those of all other simuliid genera. Molecular and morphological phylogenies both place *Cnephia* as the sister group of the most derived lineage of North American simuliids, including *Ectemnia*, *Metacnephia*, and *Simulium* (Adler et al. 2004, Moulton 2003). In the only previous

study that examined relationships within *Cnephia*, Procunier (1982) used rearrangement data to construct a "chromosome phylogeny" of the four North American species plus the Palearctic species *C. pallipes* (as *Cnephia lapponica*). Strictly speaking, diagrams of chromosomal relationship are not phylogenies because they are derived from a hypothetical "standard" that is based on its "centrality" (i.e., for any given chromosome arm, the standard sequence occurs in a number of related taxa and also gives rise to the largest number of independent derivatives) (Rothfels 1979). Accordingly, such diagrams are more appropriately described as unrooted "cytodendrograms" (Adler et al. 2004). Cytodendrograms can represent phylogenies if they are rooted using other criteria, such as information derived from morphological or molecular data.

While COI sequence data are typically analyzed using a variety of tree construction methods, the resulting barcode-based trees should not necessarily be considered phylogenetic trees (Hajibabaei et al. 2006, 2007). The properties of the mitochondrial COI gene that make it effective as a "global standard" for species recognition (i.e., rapidly evolving and maternally inherited) render it unreliable for untangling deeper-level branches on a tree. Nonetheless, a barcoding tree has potential to reflect true phylogenetic relationships if the species in question are relatively closely related. With this caveat in mind, it is instructive to compare Procunier's (1982) cytodendrogram with the topology of our MP, NJ, and Bayesian trees (Figs. 2 and 3). The latter two trees both position the Palearctic species *C. pallipes* as the plesiomorphic sister taxon of the four North American species. If *C. pallipes* (as *C. lapponica*) is used to root Procunier's (1982) cytodendrogram, the resulting topology is essentially identical to the relationships suggested by our analyses of the COI barcoding gene; *C. ornithophilia* is sister group of the other three North American species. Relationships among these latter species are unresolved cytologically, which reflects the uncertainty in our molecular analyses.

If the relationships suggested earlier are upheld following analyses of additional, more conserved genes, then there are several important evolutionary implications. Evidence suggests that *Cnephia* originated in the Palearctic Region, with subsequent dispersal to (and diversification in) the Nearctic Region. Three of the four species that evolved in North America are endemic to the Nearctic Region, whereas one species (*C. eremites*) had populations that dispersed back to the Palearctic Region, now representing the only Holarctic member of the genus.

Special comment is warranted about the evolution of feeding habits in *Cnephia.* As previously noted, the four North American species express the full range of feeding habits exhibited by female black flies. *C. eremites* and *C. dacotensis* are obligately autogenous (nonblood feeders), whereas *C. pecuarum* is mammalophilic and *C. ornithophilia* is ornithophilic. All but one *Cnephia* species for which females are known have bifid (as opposed to untoothed) tarsal claws, which is typical of black flies that blood-feed on birds (Crosskey 1990, Malmqvist et al. 2004). In fact, ornithophily is considered to be in the groundplan of the tribe Simuliini—the lineage to which *Cnephia* is assigned (Adler et al. 2004, Currie and Grimaldi 2000). Given that the common ancestor of *Cnephia* was likely ornithophilic, then autogeny and mammalophily must be secondarily derived states. Autogenous black flies occur most frequently at high altitudes or high latitudes, as is the case for the arctic-adapted species *C. eremites.* In contrast, *C. dacotensis*—the only *Cnephia* species to feature an untoothed claw—is among the few examples of a temperate-adapted autogenous species. At the other end of the feeding spectrum, *C. pecuarum* is the only major mammal feeder in North America that possesses a bifid (i.e., an ornithophilic-type) tarsal claw. While examples of host switching are known in other genera of black flies, most can be easily characterized as predominantly obligately autogenous, ornithophilic, or mammalophilic. *Cnephia* is exceptional in the catholic feeding habits exhibited by its members. Despite marked differences in the feeding habits of *Cnephia*, and all the morphological, behavioral, and physiological consequences that such differences incur (cf. Crosskey 1990), there is remarkably little reflection of such differences in the chromosomal and molecular data sets.

Implications for Biological Control. The potential significance of DNA barcoding for vector and parasite identification has long been recognized (Besansky et al. 2003). The approach has already proved successful for distinguishing many species in the most pestiferous families of biting flies, including the Simuliidae (Pramual and Kuvangkadilok 2012, Rivera and Currie 2009), Culicidae (Cywinska et al. 2006, Kumar et al. 2007, Wang et al. 2012), Ceratopogonidae (Lassen et al. 2012), and Tabanidae (Cywinska et al. 2010). Once DNA libraries have been generated for the majority of pest and vector species, it will cost a few dollars and take a few hours to identify any specimen. The implications for biological control are obvious. Large numbers of specimens can be identified quickly and easily to species level—regardless of life stage, gender, or the condition of specimens. In the case of *Cnephia*, COI barcoding offers the potential to identify breeding sites well before adults are ready to emerge, providing the time needed to implement optimal control strategies.

Acknowledgments

We thank the following individuals for providing specimens for study: P. Adler, M. Colbo, D. Craig, F. Hunter, C. LaMere, W. Lamp, J. Robinson, A. Sharp, and M. Warfel. We thank E. Gray for insights about the biological control of black flies, and D. C. Winter for assistance with Microsoft Excel scripts. This research was supported, in part, through funding to the Canadian Barcode of Life Network from Genome Canada (through the Ontario Genomics Institute), Natural Sciences and Engineering Research Council of Canada (NSERC), and other sponsors listed at www.BOLNET.ca. Additional funding was provided by an NSERC

Discovery Grant to D.C.C., and an NSERC Canada Graduate Scholarship to I.M.C.

References Cited

Adler, P. H., and R. W. Crosskey. 2012. World blackflies (Diptera: Simuliidae): a comprehensive revision of the taxonomic and geographical inventory (2012). (http://entweb.clemson.edu/biomia/pdfs).

Adler, P. H., and J. W. McCreadie. 2002. Black flies (Simuliidae), pp. 185–202. *In* G. Mullen and L. Durden (eds.), Medical and Veterinary Entomology. Academic, San Diego, CA.

Adler, P. H., D. C. Currie, and D. M. Wood. 2004. The black flies (simuliidae) of North America. Cornell University Press, Ithaca, NY.

Adler, P. H., D. Roach, W. K. Reeves, J. P. Flanagan, M. E. Morrow, and J. E. Toepfer. 2007. Attacks on the endangered Attwater's Prairie-Chicken (*Tympanuchus cupido attwateri*) by black flies (Diptera: Simuliidae) infected with an avian blood parasite. J. Vector Ecol. 32: 309–312.

Besansky, N. J., D. W. Severson, and M. T. Ferdig. 2003. DNA barcoding of parasites and invertebrate disease vectors: what you don't know can hurt you. Trends Parasitol. 19: 545–546.

Clary, D. O., and D. R. Wolstenholme. 1985. The mitochondrial DNA molecule of *Drosophila yakuba:* nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22: 252–271.

Crosskey, R. W. 1990. The natural history of blackflies. Wiley Ltd., Chichester, England.

Currie, D. C., and D. Grimaldi. 2000. A new black fly (Diptera: Simuliidae) genus from mid Cretaceous (Turonian) amber of New Jersey, pp. 473–485. *In* D. Grimaldi (eds.), Studies on Fossils in Amber, With Particular Reference to the Cretaceous of New Jersey. Backhuys Publishers, Leiden, The Netherlands.

Cywinska, A., F. F. Hunter, and P. D. Hebert. 2006. Identifying Canadian mosquito species through DNA barcodes. Med. Vet. Entomol. 20: 413–424.

Cywinska, A., M. A. Hannan, P. G. Kevan, R. W. Roughley, M. Iranpour, and F. F. Hunter. 2010. Evaluation of DNA barcoding and identification of new haplomorphs in Canadian deerflies and horseflies. Med. Vet. Entomol. 24: 382–410.

Day, J. C., M. Mustapha, and R. J. Post. 2010. The subgenus *Eusimulium* (Diptera: Simuliidae: *Simulium*) in Britain. Aquat. Insects. 32: 281–292.

Fallis, A. M., and G. F. Bennett. 1961. Sporogony of *Leucocytozoon* and *Haemoproteus* in simuliids and ceratopogonids and a revised classification of the Haemosporidiida. Can. J. Zool. 39: 215–228.

Fallis, A. M., and G. F. Bennett. 1962. Observations on the sporogony of *Leucocytozoon mirandae, L. bonasae,* and *L. fringillinarum* (Sporozoa: Leucocytozoidae). Can. J. Zool. 40: 39–400.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3: 294–299.

Fujita, M. K., A. D. Leaché, F. T. Burbrink, J. A. McGuire, and C. Moritz. 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27: 480–488.

Funk, D. J., and K. E. Omland. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34: 397–423.

Hajibabaei, M., G.A.C. Singer, and D. A. Hickey. 2006. Benchmarking DNA barcodes: an assessment using available primate sequences. Genome 49: 851–854.

Hajibabaei, M., G.A.C. Singer, E. L. Clare, and P.D.N. Hebert. 2007. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends Genet. 23: 167–172.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor, and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41: 95–98.

Hebert, P.D.N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003a. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. Bio. Sci. 270: 313–321.

Hebert, P.D.N., S. Ratnasingham, and J. R. deWaard. 2003b. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proc. R. Soc. Lond. Bio. Sci. 270: S96–S99.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17: 754–755.

Hunter, S. J., T. I. Goodall, K. A. Walsh, R. Owen, and J. C. Day. 2008. Nondestructive DNA extraction from blackflies (Diptera: Simuliidae): retaining voucher specimens for DNA barcoding projects. Mol. Ecol. Resour. 8: 56–61.

Ilmonen, J., P. H. Adler, B. Malmqvist, and A. Cywinska. 2009. The *Simulium vernum* group (Diptera: Simuliidae) in Europe: multiple character sets for assessing species status. Zool. J. Linn. Soc. 156: 847–863.

Khan, R. A., and A. M. Fallis. 1970. Life cycles of *Leucocytozoon dubreuili* Mathis and Leger, 1911 and *L. fringillinarum* Woodcock, 1910 (Haemosporidia: Leucocytozoidae). J. Protozool. 17: 642–658.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111–120.

Kumar, N. P., A. R. Rajavel, R. Natarajan, and P. Jambulingam. 2007. DNA barcodes can distinguish species of Indian mosquitoes (Diptera: Culicidae). J. Med. Entomol. 44: 1–7.

Lassen, S. B., S. A. Nielsen, H. Skovgård, and M. Kristensen. 2012. Molecular differentiation of *Culicoides* biting midges (Diptera: Ceratopogonidae) from the subgenus *Culicoides* Latreille in Denmark. Parasitol. Res. 110: 1765–1771.

Malmqvist, B., D. Strasevicius, O. Hellgren, P. H. Adler, and S. Bensch. 2004. Vertebrate host specificity of wild-caught blackflies revealed by mitochondrial DNA in blood. Proc. R. Soc. Lond. Bio. Sci. 271: S152–S155.

Molloy, D., R. Gaugler, and H. Jamnback. 1981. Factors influencing efficacy of *Bacillus thuringiensis var. israelensis* as a biological control agent of black fly larvae. J. Econ. Entomol. 74: 61–64.

Moulton, J. K. 2003. Can the current molecular arsenal adequately track rapid divergence events within Simuliidae (Diptera)? Mol. Phylogenet. Evol. 27: 45–57.

Nardi, F., A. Carapelli, P. P. Fanciulli, R. Dallai, and F. Frati. 2001. The complete mitochondrial DNA sequence of the basal Hexapod *Tetrodontophora bielanensis:* evidence for heteroplasmy and tRNA translocations. Mol. Biol. Evol. 18: 1293–1304.

Nylander, J.A.A. 2004. MrModeltest v2: program distributed by the author. Evolutionary Biology Centre, Uppsala University, Sweden.

Posada, D., and T. R. Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike

Information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53: 793–808.

Pramual, P., and C. Kuvangkadilok. 2012. Integrated cytogenetic, ecological, and DNA barcode study reveals cryptic diversity in *Simulium* (*Gomphostilbia*) *angulistylum* (Diptera: Simuliidae). Genome 55: 447–458.

Pramual, P., K. Wongpakam, and P. H. Adler. 2011. Cryptic biodiversity and phylogenetic relationships revealed by DNA barcoding of oriental black flies in the subgenus *Gomphostilbia* (Diptera: Simuliidae). Genome 54: 1–9.

Procunier, W. S. 1982. A cytological study of species in *Cnephia* s. str. (Diptera: Simuliidae). Can. J. Zool. 60: 2866–2878.

Riley, C. V. 1887. Report of the entomologist. U.S. Department of Agriculture Report 1886, pp. 459–592.

Rivera, J., and D. C. Currie. 2009. Identification of Nearctic black flies using DNA barcodes (Diptera: Simuliidae). Mol. Ecol. Resour. 9: 224–236.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Rothfels, K. H. 1979. Cytotaxonomy of black flies (Simuliidae). Annu. Rev. Entomol. 24: 507–537.

Saitou, N., and M. Nei. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406–425.

Swofford, D. L. 2003. PAUP*: phylogenetic analysis using parsimony (*and other methods) version 4. Sinauer, Sunderland, MA.

Wang, G., C. Li, X. Guo, D. Xing, Y. Dong, Z. Wang, Y. Zhang, M. Liu, Z. Zheng, H. Zhang, et al. 2012. Identifying the main mosquito species in China based on DNA barcoding. PLoS ONE 7: 1–11.

Received 27 March 2013; accepted 28 August 2013.